Finitely Generated Modules over a P.I.D.

Let R be a P.I.D. (say $\mathbb{Z}$ or $\mathbb{K}[x]$, $\mathbb{K}$ field) and $I \subseteq R$ an ideal. So $I = (a)$ for some $a \in R$.

Let M be a finitely generated R-module, say $M = \langle m_1, m_2, \dots, m_n \rangle$.

Consider the R-module homomorphism $\varphi : R^n \longrightarrow M$ given by $\varphi(e_i) = m_i$, $i = 1, \dots, n$, $e_i = (0, \dots, \underset{i^{th} \text{ place}}{1}, \dots 0)$.

Since φ is onto we see that

$$M = \operatorname{Im}\varphi \simeq R^n / \ker\varphi.$$

__Example:__ $R = \mathbb{Z}$, say $M = \mathbb{Z}\times\mathbb{Z} / \langle (5,2), (3,4) \rangle$,

$\varphi : R^2 \longrightarrow M$, $\varphi(1,0) = \overline{(1,0)}$ and $\varphi(0,1) = \overline{(0,1)}$.

We'll see that $M = \mathbb{Z}\times\mathbb{Z} / \langle (5,2), (3,4) \rangle$

$$\simeq \mathbb{Z}_{14}.$$

Note that $\det\begin{pmatrix} 5 & 2 \\ 3 & 4 \end{pmatrix} = 14$!

Indeed, $\begin{pmatrix} 5 & 2 \\ 3 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} -1 & -6 \\ 3 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} -1 & -6 \\ 0 & -14 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 14 \end{pmatrix}$

and $M \simeq \mathbb{Z}^2 / \langle (5,2), (3,4) \rangle = \dfrac{\langle (1,0), (0,1) \rangle}{\langle (5,2), (3,4) \rangle} = \dfrac{\langle (1,0), (0,1) \rangle}{\langle (1,0), (0,14) \rangle} \simeq \mathbb{Z}_{14}.$

Smith Normal Form:

Let R be a P.I.D. and $M = \langle m_1, \dots, m_n \rangle$ finitely generated R-module.

$$\varphi : R^n \to M, \quad \varphi(e_i) = m_i, \quad e_i = (0, \dots, 1, \dots, 0), \; i = 1, \dots, n.$$

$M \simeq R^n / \ker\varphi$. $\quad \ker\varphi = \langle f_i \mid i \in \Lambda \rangle$

Let $f_i = a_{i1} e_1 + \dots + a_{in} e_n$ and consider the coefficient matrix (possibly infinite size!)

$$A = \begin{array}{c} \begin{array}{cccc} e_1 & e_2 & & e_n \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{i1} & a_{i2} & & a_{in} \\ \vdots & \vdots & & \vdots \end{pmatrix} \end{array}$$

Consider the following row and column operations that will correspond to "base change operations" for R^n and the submodule $\ker\varphi$.

$C_1)$ $e_j \to e_j + \lambda e_i$, $\lambda \in R$

$$\langle e_1, \dots, e_j, \dots, e_n \rangle \longmapsto \langle e_1, \dots, e_j + \lambda e_i, \dots, e_n \rangle$$

$C_2)$ $e_i \leftrightarrow e_j$, $\langle e_1, \dots, e_i, \dots, e_j, \dots, e_n \rangle \mapsto \langle e_1, \dots, e_j, \dots, e_i, \dots, e_n \rangle$

$C_3)$ $e_i \to u e_i$, $u \in R$ is a unit

$$\langle e_1, \dots, e_i, \dots, e_n \rangle \longmapsto \langle e_1, \dots, u e_i, \dots, e_n \rangle.$$

In terms of the coefficient matrix these are just the following column operations:

$C_1)$ $f_k = a_{k1}e_1 + \cdots + a_{ki}e_i + \cdots + a_{kj}e_j + \cdots + a_{kn}e_n$

$$= a_{k1}e_1 + \cdots + (a_{ki} - \lambda a_{kj})e_i + \cdots + a_{kj}(e_j + \lambda e_i) + \cdots + a_{kn}e_n$$

So C_1 is the column operation which replaces the j^{th} column with the j^{th} column $+ \lambda\, i^{th}$ column.

$C_2)$ $e_i \longleftrightarrow e_j$ interchanges the i^{th} and j^{th} columns of the matrix A.

$C_3)$ $ue_i \longrightarrow e_i$ just multiplies the i^{th} column by the unit $u \in R$

On the other hand, the corresponding operations on the generating set $\{f_i\}$ correspond row operations on A.

$R_1)$ $f_j \longrightarrow f_j + \lambda f_i$

$\langle f_1, \ldots, f_i, \ldots, f_j, \ldots \rangle \longrightarrow (f_1, \ldots, f_i, \ldots, f_j + \lambda f_i, \ldots)$

$f_j = a_{j1}e_1 + \cdots + a_{jn}e_n$

$f_j + \lambda f_i = (a_{j1} + \lambda a_{i1})e_1 + \cdots + (a_{jn} + \lambda a_{in})e_n$

So, the j^{th} row is replaced by j^{th} row $+ \lambda\, i^{th}$ row.

$R_2)$ $f_i \leftrightarrow f_j$ just replaces the i^{th} and j^{th} rows.
$R_3)$ $f_i \longrightarrow uf_i$, just multiplies the i^{th} row by $u \in R$.

Since R is a P.I.D. it is a Unique Factorization Domain. So any element of R has a unique factorization into prime elements.

Example 1) $R = \mathbb{Z}$, $n \in \mathbb{Z}$, $n = p_1^{r_1} \cdots p_k^{r_k}$, p_i prime

2) $R = \mathbb{K}[x]$, $f \in R$, $f = p_1^{r_1}(x) \cdots p_k^{r_k}(x)$, $p_i \in K[x]$ irreducible polynomials.

Let $o(x)$ denote the number of prime elements in the prime decomposition of x.

For example, if $R = \mathbb{Z}$, $o(1) = 0$, $o(5) = 1$, $o(35) = 2$, $o(16) = 4$.

— o —

Back to Smith Normal Form:

$M = \langle m_1, \ldots, m_n \rangle$ R-module, $\varphi : R^n \to M$, $\varphi(e_i) = m_i$.
Let $L = \ker \varphi$, so that $M \cong R^n / L$.

Say $L = \langle f_j \rangle$, $j \in \Lambda$.

$$A = \begin{array}{c} \begin{array}{cccc} e_1 & e_2 & \cdots & e_n \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \end{pmatrix} \end{array} \begin{array}{c} f_1 \\ f_2 \\ \vdots \\ f_j \\ \end{array}$$

If A is not the zero matrix by replacing rows and columns we may assume that $a_{11} \neq 0$. If $a_{11} | a_{k1}$ then by the row operation $R_k \to R_k - \lambda R_1$, where $\lambda a_{11} = a_{k1}$, we can make $a_{k1} = 0$. If $a_{11} \nmid a_{k1}$ but $a_{k1} | a_{11}$ then replace $R_1 \leftrightarrow R_k$ so that we obtain $a_{11} | a_{k1}$. Then we can again make $a_{k1} = 0$ as before. If $a_{11} \nmid a_{k1}$ and $a_{k1} \nmid a_{11}$ let $c = (a_{11}, a_{k1})$. Then $o(c) < o(a_{11})$ and $o(c) < o(a_{k1})$. Say $c = \sigma a_{11} + \tau a_{k1}$, $\sigma, \tau \in R$. Then

$$\sigma \frac{a_{11}}{c} + \tau \frac{a_{k1}}{c} = 1.$$

Now the invertible operation

$$e_1, e_2 \longmapsto e_1' = \sigma e_1 + \tau e_2, \; e_2' = -\frac{a_{k1}}{c} e_1 + \frac{a_{11}}{c} e_2,$$

which corresponds to the matrix multiplication

$$(*) \quad \begin{pmatrix} \sigma & \tau & 0 & \cdots & 0 \\ -\frac{a_{k1}}{c} & \frac{a_{11}}{c} & 0 & & 0 \\ 0 & 0 & 1 & & \\ \vdots & \vdots & & \ddots & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & & \end{pmatrix} = \begin{pmatrix} c & * & * \\ 0 & * & * \\ & * & \end{pmatrix}$$

(here we took $k=2$, for simplicity. This matrix operation may not be an elementary row/column operation!)

Performing this operations finitely many times we may bring the matrix A to the form

$$\begin{pmatrix} c & * & * \\ 0 & * & * \\ 0 & & \\ \vdots & * & \\ 0 & & \end{pmatrix}.$$

Then we perform column operation to bring the matrix to a form $\begin{pmatrix} d & 0 & - & 0 \\ * & & & \\ * & & & \\ * & & & \end{pmatrix}$.

Note that we will have $o(d) \leq o(c)$. Then we repeat these operations and obtain

$$\begin{pmatrix} * & * & * \\ 0 & & \\ \vdots & * & \\ 0 & & \end{pmatrix}$$

again. Since $o(c)$ will decrease each time we'll finally get the matrix of the form

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}.$$

Now we repeat similar operation for the submatrix B and so on. Finally we'll obtain a matrix of the form

$$\begin{pmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{pmatrix}.$$

Indeed by performing more operations we may assume that

$d_1 | d_2,\ d_2 | d_3, \dots, d_{n-1} | d_n.$

Now we have $R^n = \langle e_1, \dots, e_n \rangle \supseteq \langle f_j \mid j \in \Lambda \rangle = L$

$$R^n = \langle e_1', \dots, e_n' \rangle \supseteq \langle f_1', \dots, f_n' \rangle = L.$$

This implies that $f_i' = d_i e_i'$, for all $i = 1, \dots, n$.

Hence, $M \simeq R^n/L = \dfrac{\langle e_1', \dots, e_n' \rangle}{\langle f_1', \dots, f_n' \rangle} = \dfrac{\langle e_1', \dots, e_n' \rangle}{\langle d_1 e_1', \dots, d_n e_n' \rangle}$

$$\simeq R/d_1R \oplus R/d_2R \oplus \cdots \oplus R/d_nR.$$

<u>Remark</u>: Let R be a Euclidean domain, say $R = \mathbb{Z}$ or $\mathbb{K}[x]$. Then there is a function $n: R \to \mathbb{N}$ so that for any $x, y \in R$, with $y \neq 0$, we have $x = qy + r$ for some $q \in R$ and $r \in R$ with $n(r) < n(y)$. This case all invertible operations we apply to the matrix A can be chosen as row or column operations. (See the matrix operation $*$)

Theorem: The elements d_1, d_2, d_n satisfying $d_1 | d_2, \ldots, d_{n-1} | d_n$ are uniquely determined up to multiplication by units. They will be called invariant factors of the R-module M.

Proof: Δ_1 = g.c.d. of all 1×1-minors of A
= g.c.d. of all 1×1-minors of D
$= d_1$

$$A = (a_{ij}) \longrightarrow \cdots \longrightarrow \begin{pmatrix} d_1 & & \\ & d_2 & \\ & & \ddots \\ & & & d_n \end{pmatrix} = D.$$

This statement is true because our operations does not change the the g.c.d. of the minors of A.

Similarly, let Δ_2 = g.c.d. of all 2×2 minors of A
= g.c.d. of all 2×2 minors of D, and
so on until
Δ_n = determinant of A
= determinant of D.

In particular, $d_1 = \Delta_1$, $d_2 = \frac{\Delta_2}{\Delta_1}, \ldots, d_n = \frac{\Delta_n}{\Delta_{n-1}}$.

This finishes the proof. ∎

Example: Let $R = \mathbb{Z}$, then any finitely generated R-module is nothing but an abelian group. Then by the above considerations
$$M \simeq \mathbb{Z}/d_1\mathbb{Z} \oplus \mathbb{Z}/d_2\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/d_n\mathbb{Z},$$ where

$d_1 | d_2, \dots, d_{n-1} | d_n$. Note that if some $d_i = 0$ then $\mathbb{Z}/d_i\mathbb{Z} = \mathbb{Z}/(0) \simeq \mathbb{Z}$.

So we get $M \simeq \mathbb{Z}_{d_1} \oplus \cdots \oplus \mathbb{Z}_{d_k} \oplus \mathbb{Z} \oplus \cdots \oplus \mathbb{Z}$.

For instance if $(d_1, d_2, \dots, d_5) = (2, 6, 18, 0, 0)$ then

$M \simeq \mathbb{Z}_2 \oplus \mathbb{Z}_6 \oplus \mathbb{Z}_{18} \oplus \mathbb{Z} \oplus \mathbb{Z}$.

Canonical Forms of Operators on Finite dim'l Vector Spaces:

$\mathbb{K}$ a field, $V = \mathbb{K}^n$ and $T : V \to V$ an operator.

$R = \mathbb{K}[X]$ is a Euclidean domain and hence a P.I.D.

We may regard $V = \mathbb{K}^n$ as an $R = \mathbb{K}[X]$-module as follows:

$R \times V \longrightarrow V$, $(p(x), v) \longmapsto p(x) \cdot v = p(T)(v)$

<u>Example</u>: Let $p(x) = 3x^2 - 5x + 4$, then $p(T) = 3T^2 - 5T + 4I$.

So, for any $v \in V$, $p(x) \cdot v = p(T)(v)$

$$= 3T^2(v) - 5T(v) + 4v$$

Define the following R-module homomorphism

$\varphi : R^n \to V$, given by $\varphi(e_i^R) = e_i$, $i = 1, \dots, n$, where

$e_i \in V = \mathbb{K}^n$, $e_i = (0, \dots, 1, \dots 0)$ and $e_i^R \in R^n$, $e_i^R = (0, \dots, 1, \dots 0)$

By its definition φ is clearly onto: $V = \operatorname{Im} \varphi$.

Hence, $V = \operatorname{Im} \varphi \simeq R^n / \ker \varphi$.

Note that $\varphi(p_1(x), \dots, p_n(x)) = \varphi\left(\sum_{i=1}^{n} p_i(x) e_i^R\right)$

$$= \sum_{i=1}^{n} p_i(x) \left(\varphi(e_i^R)\right)$$

$$\Rightarrow \varphi(p_1(x), \dots, p_n(x)) = \sum_{i=1}^{n} p_i(x)(e_i)$$
$$= \sum_{i=1}^{n} p_i(T)(e_i).$$

~ o ~ o ~ o ~

For the basis $\mathcal{B} = \{e_1, \dots, e_n\}$ of $V = K^n$, let

$C = [T]_{\mathcal{B}}^{\mathcal{B}}$. Hence, $Te_i = \sum_{j=1}^{n} c_{ji}\, e_j$.

This implies that $-c_{1i}e_1 - \dots + (T - c_{ii})e_i - \dots - c_{ni}e_n = 0$.
In other words, $(-c_{1i}, -c_{2i}, \dots, T - c_{ii}, \dots, -c_{ni}) \in \ker\varphi$.

Note that $(-c_{1i}, -c_{2i}, \dots, T - c_{ii}, \dots, -c_{ni})^T$ is the i^{th} column of

$$xI - C = \begin{bmatrix} x - c_{11} & -c_{12} & \cdots & -c_{1i} & \cdots & -c_{1n} \\ -c_{21} & x - c_{22} & & -c_{2i} & & -c_{2n} \\ \vdots & \vdots & & x - c_{ii} & & \vdots \\ -c_{n1} & -c_{n2} & & -c_{ni} & & x - c_{nn} \end{bmatrix}.$$

<u>Lemma</u>: The submodule $\ker\varphi$ is generated by the columns of $xI - C$.

<u>Proof</u>: We have already seen that $\ker\varphi$ contains the columns of $xI - C$. For the other direction we proceed as follows.

Let D be the Smith Normal Form of $xI - C$.

$$xI - C \rightsquigarrow D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix} \quad d_i \mid d_{i+1}, \; i = 1, \dots, n-1.$$

Let U be the submodule generated by the columns of $xI - C$. Then U is generated by the columns of D. (Recall that in our discussion of Smith Normal Form f_i's were the rows of the matrix A, not the columns. However, this does not cause any problems! Why?)
Now the above observation implies

$$R^n/U \simeq R^n/\langle d_1 e_1^R, \dots, d_n e_n^R \rangle \simeq R/d_1R \oplus \dots \oplus R/d_nR$$

Note that $d_1 \cdot d_2 \cdots d_n = \det D = \det(xI - C)u$, for some $u \in \mathbb{K} \setminus \{0\}$, unit of $R = K[x]$.

$\deg(d_1 \cdots d_n) = \deg(xI - C) = n \Rightarrow \sum_{i=1}^{n} \deg(d_i) = n$, where each $d_i = x^k + a_{k-1}^i x^{k-1} + \dots + a_1^i x + a_0^i$, for some $a_j^i \in \mathbb{K}$.

Thus, $R/d_iR = \mathbb{K}[x]/(d_i) \simeq \langle 1, x, \dots, x^{k-1} \rangle$ as R-modules.

The isomorphism is also an $\mathbb{K}$-vector space isomorphism. Note that $\{1, x, \dots, x^{k-1}\}$ is a $\mathbb{K}$-basis for R/d_iR. Hence, $\dim_{\mathbb{K}} R/d_iR = k$. Hence, $\mathbb{K}$-vector spaces,

$$R/U \simeq R/d_1R \oplus \dots \oplus R/d_nR \simeq \mathbb{K}^{\deg(d_1)} \oplus \dots \oplus \mathbb{K}^{\deg(d_n)}$$
$$\simeq \mathbb{K}^n$$

Since $U \subseteq \ker \varphi$ we have $\mathbb{K}^n \simeq R^n/U \xrightarrow{\phi} R^n/\ker\varphi \simeq \mathbb{K}^n$ where ϕ is induced by the identity map $R^n \to R^n$. Clearly, ϕ is onto. Since both vector spaces are isomorphic to $\mathbb{K}^n$ they have the same dimension n and thus $\bar{\phi}$ is an isomorphism. This implies $(0) = \ker \bar{\phi} = \ker \varphi/_U \Rightarrow U = \ker \varphi$, which finishes the proof. ■

<u>Corollary</u> $V \simeq \mathbb{K}^n \simeq R^n/_{\ker\varphi} \overset{\phi}{\simeq} R^n/U \simeq R/_{d_1R} \oplus \cdots \oplus R/_{d_nR}$

Note that the above isomorphism are both R-module isomorphisms and $\mathbb{K}$-vector space isomorphisms.

$$V = \mathbb{K}^n \longrightarrow \mathbb{K}[x]/_{d_1\mathbb{K}[x]} \oplus \cdots \oplus \mathbb{K}[x]/_{d_n\mathbb{K}[x]}$$

$$\begin{tikzcd} V \arrow[r, "\phi"] \arrow[d, "T"'] & R^n/U \arrow[d, "x"] \\ V \arrow[r, "\phi"] & R^n/U \end{tikzcd} \qquad \begin{tikzcd} v \arrow[r, mapsto] \arrow[d, mapsto] & {[p(x)]} \arrow[d, mapsto] \\ T(v) \arrow[r, mapsto] & {[x\, p(x)]} \end{tikzcd}$$

<u>Observation:</u>

1) <u>Cayley Hamilton Theorem:</u>

We know that $f(x) = \det(xI - C) = d_1 d_2 \cdots d_n$.

Thus, $f(T)(v) \longrightarrow [f(x)p(x)] = 0$ in

$\mathbb{K}[x]/_{d_1} \oplus \cdots \oplus \mathbb{K}[x]/_{d_n}$.

Hence, $f(T)(v) = 0$, for all $v \in V$. This implies

$f(T)=0$ is the zero operator. So we have just proved

<u>Theorem</u> (Cayley–Hamilton Thm)

For any operator $T: V \to V$, $(V \cong \mathbb{K}^n)$ we have $f(T)=0$, where $f(x)=\det(xI-T)$, the characteristic polynomial of T.

2) Since each $d_i \mid d_n$, $i=1,\dots,n$, we see that d_n is the lowest degree polynomial in $\mathbb{K}[x]$ so that $d_n(T)=0$. Therefore, d_n will be called the minimal polynomial of $T: V \to V$.

3) $V = \bigoplus_{i=1}^{n} \mathbb{K}[x]/d_i$

Let $d_i = a_0 + a_1 x + \dots + a_{k-1}x^{k-1} + x^k$. So, $\{1, x, \dots, x^{k-1}\}$ is a basis for the subspace $\mathbb{K}[x]/d_i$.

Since $x \cdot 1 = x$, $x \cdot x = x^2, \dots, x \cdot x^{k-1} = x^k = -a_0 - a_1 x - \dots - a_{k-1}x^{k-1}$, because $d_i = a_0 + \dots + x^k = 0$ in $\mathbb{K}[x]/d_i$.

It follows that T has matrix representation

$$[T] = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_n \end{bmatrix},$$ where A_i is the representation of $T|_{\mathbb{K}[x]/d_i}$, and $A_i = \begin{bmatrix} 0 & \cdots & \cdots & 0 & -a_0 \\ 1 & & & & -a_1 \\ \vdots & \ddots & & & \vdots \\ 0 & \cdots & & 1 & -a_{k-1} \end{bmatrix}$

This is called the rational form of T.

Indeed, we can do more. Say $d(x) = p_1^{r_1} p_2^{r_2} \cdots p_k^{r_k}$, where each p_i is irreducible polynomial in $\mathbb{K}[x]$.
Then
$$\mathbb{K}[x]/_d = \mathbb{K}[x]/_{p_1^{r_1} \cdots p_k^{r_k}} \simeq \mathbb{K}[x]/_{p_1^{r_1}} \oplus \cdots \oplus \mathbb{K}[x]/_{p_k^{r_k}}.$$

5) Jordan Canonical Form:

Now assume that the characteristic polynomial $f(x) = \det(xI - T)$ is a product of linear terms. Hence, each d_i is a product of linear terms. Say, $d_i = (x-\lambda_1)^{r_1} \cdots (x-\lambda_k)^{r_k}$. Note that if $K = \mathbb{C}$ this is always the case.

Now $\mathbb{K}[x]/_{d_i} \simeq \mathbb{K}[x]/_{(x-\lambda_1)^{r_1}} \oplus \cdots \oplus \mathbb{K}[x]/_{(x-\lambda_k)^{r_k}}$.

The subspace $\mathbb{K}[x]/_{(x-\lambda)^r}$ has basis

$\{v_0 = 1, v_1 = x-\lambda, v_2 = (x-\lambda)^2, \ldots, v_{r-1} = (x-\lambda)^{r-1}\}$.

Also note that

$x \cdot v_0 = x \cdot 1 = x - \lambda + \lambda \cdot 1 = \lambda v_0 + 1 \cdot v_1$

$x\, v_1 = x(x-\lambda) = x^2 - \lambda x = (x-\lambda)^2 + \lambda(x-\lambda) = \lambda v_1 + v_2$

$\vdots$

$x\, v_{r-2} = x(x-\lambda)^{r-2} = (x-\lambda)^{r-1} + \lambda(x-\lambda)^{r-2} = \lambda v_{r-2} + v_{r-1}$

$x\, v_{r-1} = x(x-\lambda)^{r-1} = (x-r)^{\lambda} + \lambda(x-\lambda)^{r-1} = 0 + \lambda v_{k-1} = \lambda v_{k-1}$.

Hence, the matrix representation of T in this subbasis is

$$J_{\lambda,r} = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 1 & \lambda & & \vdots \\ & \ddots & \ddots & 0 \\ 0 & & 1 & \lambda \end{bmatrix}_{r \times r}$$

This is called a Jordan block. So, T has a matrix representation consisting of Jordan blocks:

$$[T] = \begin{bmatrix} J_{\lambda_1, r_1} & & \\ & \ddots & \\ & & J_{\lambda_{n_0}, r_{n_0}} \end{bmatrix}.$$

6) Characterization of Diagonalizibility:

Theorem: $T: V \to V$, $V \simeq K^n$, is diagonalizable if and only if the minimal polynomial $d_n(x)$ of T is a product of distinct linear factors.

Proof: Assume that the minimal polynomial $d_n = d_n(x)$ is a product of distinct linear factors:

$$d_n(x) = (x-\lambda_1)\cdots(x-\lambda_k), \quad \lambda_i \neq \lambda_j \text{ if } i \neq j$$

Since $\det(xI - T) = f(x) = d_1(x) \cdots d_n(x)$ and $d_i \mid d_{i+1}, \forall i$, we see that each $d_i(x)$ and $f(x)$ is a product of linear factors. Moreover, each $d_i(x)$ is also a product of linear factors. Also, $f(x)$ and

$d_n(x)$ have the same linear factors, where the multiplicities in $f(x)$ may not be one!

For any eigenvalue λ_i if n_i is the multiplicity of $(x-\lambda_i)$ in $f(x)$ then the solution space of $\lambda_i I - T = 0$ has dimension exactly n_i:

$$(\lambda_i I - T)v = 0 \Longleftrightarrow \begin{bmatrix} d_1(\lambda_i) & & \\ & \ddots & \\ & & d_n(\lambda_i) \end{bmatrix} v = 0, \text{ because}$$

exactly n_i many of $d_1(\lambda_i), \ldots, d_n(\lambda_i)$ are zero.

Since $\dim V = n = \deg f(x) = \sum_{i=1}^{k} n_i = \sum_{i=1}^{k} \dim W_{\lambda_i}$,

where W_{λ_i} is the eigenspace of the eigenvalue λ_i, we see that the operator T is diagonalizable.

Similarly, if T is diagonalizable then again by the formula
$n = \sum_{i=1}^{k} n_i$ we see that the solution space of $\begin{bmatrix} d_1(\lambda_i) & & \\ & \ddots & \\ & & d_n(\lambda_i) \end{bmatrix} v = 0$ is n_i so that each $d_j(x)$ consists of distinct linear factors and hence $d_n(x)$ is a product of distinct linear factors.
This finishes the proof. ∎

7) <u>Similar Matrices</u>! Let A and B be similar matrices. Hence, $B = P^{-1}AP$ for some invertible P. Then $xI - B = P^{-1}(xI - A)P$

and thus $xI-A$ and $xI-B$ have the same Smith Normal Forms.

Conversely, if $xI-A$ and $xI-B$ have the same Smith Normal Forms then A and B have the same canonical forms, say rational form C. Then $P_1^{-1}AP_1 = C = P_2^{-1}BP_2$ for some matrices P_1 and P_2. Then

$$B = P_2P_1^{-1}AP_1P_2^{-1} = P^{-1}AP,$$ where $P = P_1P_2^{-1}$. In particular, A and B are similar. ◾